A Framework for a Priori Evaluation of Multimodal User Interfaces Supporting Cooperation

Magnus Larsson, Gilles Coppin, Franck Poirier, Olivier Grisvard

A Framework for a Priori Evaluation of Multimodal User Interfaces Supporting Cooperation

Magnus Larsson
TELECOM Bretagne
Technopôle Brest-Iroise - CS 83818 - 29238 Brest Cedex 3 - France
magnus.larsson@telecom-bretagne.eu

Gilles Coppin
TELECOM Bretagne
Technopôle Brest-Iroise - CS 83818 - 29238 Brest Cedex 3 - France
gilles.coppin@telecom-bretagne.eu

Franck Poirier
Université de Bretagne-Sud
BP 573, Vannes Cedex, 56017 France
franck.poirier@univ-ubs.fr

Olivier Grisvard
TELECOM Bretagne
Technopôle Brest-Iroise - CS 83818 - 29238 Brest Cedex 3 - France
olivier.grisvard@telecom-bretagne.eu

ABSTRACT
In this short paper we will present our latest research on a new framework being developed for aiding novice designers of highly interactive, cooperative, multimodal systems to make expert decisions in choice of interaction modalities depending on the type of activity and its cooperative nature. Our research is conducted within the field of maritime surveillance at ATOL Laboratory and it is focused on the next generation distributed work support.

Author Keywords
Analysis, *a priori* evaluation; cooperation; multimodality; system design.

ACM Classification Keywords
H.1.2: Human factors; H5.2: Evaluation/methodology

General Terms
Design; Human Factors; Management

INTRODUCTION
The computer industry is on the brink of a new era. The future is not a solitary PC, but a diverse set of smart, cooperative devices interacting not only with its end users but also with each other while fully integrated in their environment. The interaction with these systems are multimodal where the tools become extensions of the human sensor and motor systems supporting the end users' cooperative execution of actions while trying to solve problems. The computer is thus no longer a system that just determines something by mathematical means, brings order (Fr. 'Ordinateur'), handles data (Swe. 'Dator'), count information (Hun. 'Bilgisayar'), or is a machine full of knowledge (Fin. 'Tietokone'). It is rather an infrastructure for multimodal human-computer interaction and cooperation. However, are we as designers equipped to meet the rapid evolution within the computer industry? In this short paper, we suggest that we need to find a way to minimize the gap between analysis and design to be able to continue delivering optimized and satisfactory systems to our customers and end users at a reasonable price.

CHI'12, May 5–10, 2012, Austin, Texas, USA.

THE DESIGNERS' CHALLENGES OF TODAY
The vast majority of today's expert designers are still novices within the design of highly interactive, cooperative, multimodal systems. However, they are still supposed and demanded to deliver intuitive, useful systems of high quality to a reasonable cost. The technology necessary to create these systems are mere a mash-up of existing technologies. Model-driven languages, methods and tools are continuously being developed and enhanced to meet the demands of the industry on adaptive, flexible and robust [8] systems designed and developed at a low cost. One example of such a project is the pan-European ITEA2: UsiXML project which is based on the µ7 concept, i.e. multi-device, multi-platform, multi-user, multi-linguality/culturality, multi-organization, multi-context, and multi-modality. However, due to the designers' lack of experience and know-how in designing these new complex systems, and due to the intended end users' and customers' inability to clarify and articulate their cooperative and multimodal needs in a comprehensive way, the designers often face infoglut resulting in poor choices in interaction modalities. Some of the most common challenges are:

- The intuition and decision-making of the designers regarding computer supported cooperative multimodal systems are biased by previous experiences of single-user system design
- The complexity in group interactions and activities pose great challenges:
    - Group logistics of data collection
    - Number and complexity of variables
    - Validation of re-engineered group work
- It is time and money consuming to perform evaluation of multimodal cooperation even though one focus on a smaller set of activities and well

defined user groups (even for a one-user application)

- The lowest common denominator is "easily" validated for single user systems, but not for multimodal cooperative systems with a big variety of end users
- There is a disparity in activity objective and needs between who does the work (the end users) and who gets the benefit of that same cooperative work (the management team/the customer)

One way to aid the designers would be to provide a framework that can alleviate the transition from analysis to design. Today, this is a tedious time consuming work biased on deficient mental models by the designers and without any promise of quality delivered. Therefore, our intention is to help novice designers of cooperative multimodal systems to make expert decisions in choice of modality or combinations of modalities. We believe that this will not only enable the creation of new, for the end users, adequate intuitive systems supporting their cooperative work, but it will also optimize the projects' ROI. In the following paragraphs we will present an *a priori* evaluation framework being developed based our understanding of human behavior and cooperation and how multimodal interaction can be approached to solve these issues.

HUMAN BEHAVIOR AND COOPERATION

One of the strongest assets of human beings are their ability to interact with each other in quite complex ways in order to fulfill a great number of simultaneous tasks initiated from a wide variety of intentions [7]. These interactions take place within a group of people, i.e. two or more participants, who can be considered to cooperate to the extent that they 1) consider each other cognitively in interaction, 2) have a joint purpose, 3) consider each other ethically in interaction, and 4) trust each other to act according to 1-3 [1]. Novices and experts meet in different groups and teams within which they take on passive, active or expansive roles, while belonging to different communities of interest and practice at the same time. [4]. Their interactions can be collective or dispersed and they can be direct, i.e. interpersonal, or indirect, e.g. mediated by computers. Furthermore, depending on their level of involvement, one can consider them to engage in no interaction, lightweight interaction, information sharing, coordination, collaboration or cooperation. Evidently, the complexity of human interaction together with the challenges posed on the designer regarding choice of interaction modality demands a comprehensive framework to avoid infoglut in the moving from analysis to design.

Based on the work of prominent scientists during mid and late 19th century, such as Charles Robert Darwin, Gustav Theoder Fechner and Mikhaylovich Sechenov, the Russian psychologist Lev Semyonovich Vygotsky founded cultural-historical psychology, thus closing the gap between the natural sciences and the mental sciences of human behavior. He approached behavior not as a result but rather as a process in motion and in change, i.e. by studying behavior as interaction. Vygotsky's research on activities bridged the gap between the mental and the physical contexts of human behavior and consciousness [10]. Activity Theory, an evolution of Vygotsky's research, provides a basic framework for human interaction and a useful basic unit of analysis; the activity.

The Activity Theory concept deals with a set of fundamental types [5], which are:

- An object – Activities can be distinguished by their objects. It is the object and the transformation of that set object that drives the activity.
- A collective phenomenon – The activity does not take place in isolation but is always a collective phenomenon.
- A subject (agent) – The activity has a subject or a collective of subjects who understands the motive of the activity. In our research we refer to the subject as an actor or a role.
- A material environment – The activity exists in and transforms its material environment.
- A historically developing phenomenon – The activity is a process that has a shared memory.
- Contradictions – The force behind the development of an activity are contradictions.
- Actions – Participants realize an activity through conscious and purposeful actions.
- Culturally mediated relationships

These fundamental types can easily be illustrated in a diagram together with their individual relationships. Kuutti's research [5] on Activity Theory and its fundamental types has resulted in a useful framework for research on computer-supported cooperative work. Cadier [2] extended the Activity Theory framework of Kuutti to manage both negotiation and execution of cooperative work (see Figure 1 and Figure 2, below), thus enabling analysis of cooperative activity.

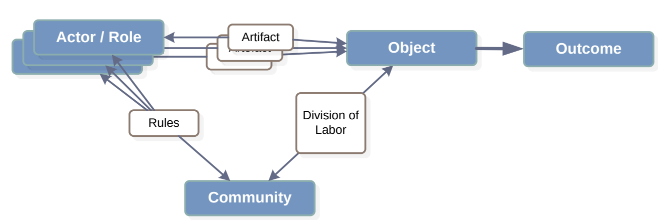

Figure 1. Basic structure of a cooperative activity [2] with its properties visualized and given a relation to each other with mediating artifacts, rules and division of labor being multimodal in their character.

The model in Figure 1, above, depicts the 'playground' of an activity, whereas the model in Figure 2, below, illustrates the actual execution process of an activity and its sub-activities/tasks and operations. Furthermore, this model also illustrates the cooperative steps of an activity where the negotiation of division of labor is the starting point, but also the result.

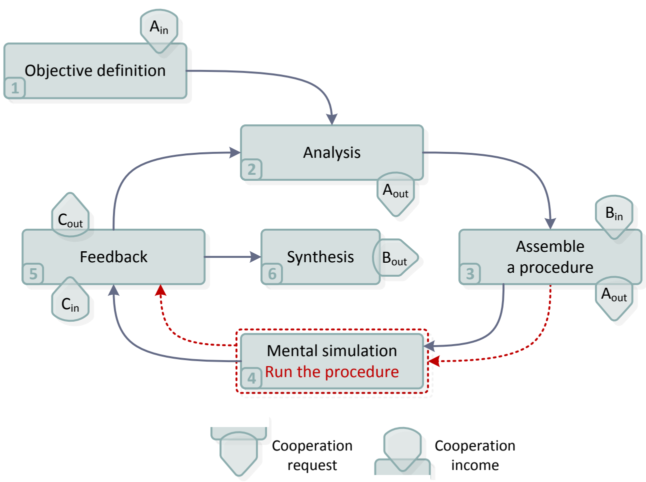

Figure 2. An iterative cooperative activity process [2] where cooperative requests are multimodal acts of negotiation and decision whilst the rest of the process is multimodal (internal as well as external) actions.

Based on this knowledge we can conclude that cooperation is heterogeneous where contradictions force activities [5], that it is culturally and contextually situated and that it makes use of internal as well as external communication [10], both verbal and non-verbal [1], to organize the same activities. These activities the users later execute with the help of mediating artifacts such as computers (see Figure 3, below). We can also conclude that the level of verbal versus non-verbal communication depends on the social context of the actor/role, which are mediated via social rules and norms and the activity's division of labor. This would suggest that no person act in isolation and that one could consider all activities, if taking into account the different levels and types of human interaction and work support, as cooperative as well as multimodal.

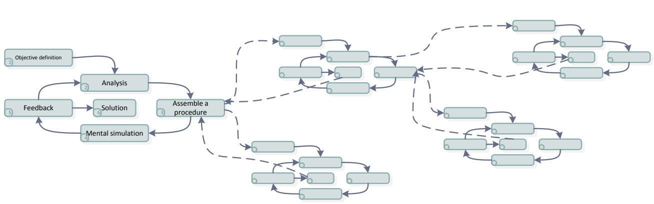

Figure 3. An example of an activity decomposition via negotiation of division of labor and the work task execution, both influenced by the social rules of the community.

Based on this understanding of human interaction and computer supported cooperative work we can look closer at what implications this has on the choice of interaction modalities and how we can develop a framework suitable for designers.

A SOUND CHOICE OF INTERACTION MODALITY

Human behavior, interaction and cooperation are multimodal by origin. Activity Theory provides, as shown above, a comprehensive high-level framework for organizing cooperative activities into manageable entities. However, in order to be able to provide any insight into preferred choice of multimodality for any specific context we need to enhance and develop it further. To be able to manage the cognitive aspects of the actors/roles in cooperation we can make use of Endsley's Situation Awareness model [3], which, in combination with our understanding of the human sensor and motor system provide a mind and body description of human capabilities (see Figure 4, below). The actor/role has been given a physical interface in his/her motor and sensor system as well as a detailed description of his/her mental capabilities. The mind and body together defines the activity process from situation awareness via decision to execution of the activity, either alone or in cooperation with other actors/roles. The actors'/roles' mental capabilities and properties are affected by the community's social and cultural rules and vice versa and the actors'/roles' use of artifacts to transform the object of interest into sought for outcome also transforms the actor/role in that same process.

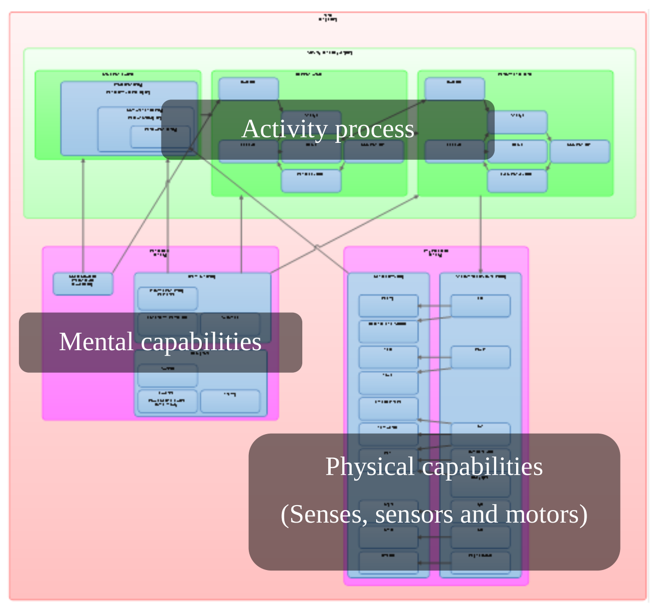

Figure 4. Our latest actor/role model together with the activity process together with the mental as well as physical capabilities based on work of [2], [3], [4] and [5] [model construction in progress]

A computer system also has a physical interface towards the outside world and an inner "mental" core based on the intentions of the system creators as well as its users' use and re-work. The physical aspects (input and output devices) together with the logical interaction language make an interaction modality which together with e.g. CARE properties can be combined into multimodal interactive systems [6] while providing *plasticity* [8] (see Figure 5, below) to correspond to the changing context.

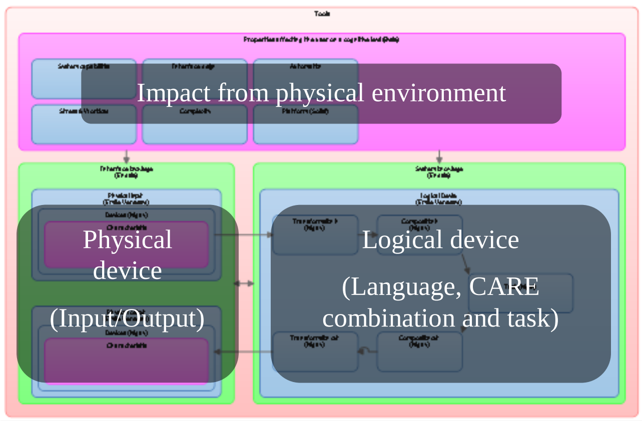

Figure 5. Our latest artifact model of the computer interface based on the work of [6], [8], [9] and [10].

The understanding that human interaction with computer systems and other humans varies depending on if the actor/role is acting alone or in cooperation with other humans or computers as well as on the context of use is crucial. By comparing the actors'/roles' mental capabilities and objectives with the logical device and its intended task support of the computer system one can evaluate the mental aspect of the activity. By comparing the physical capabilities of the actors/roles with each other and with the computer system one can find constraints as well as possibilities of interaction modalities. The actors'/roles' negotiation with the group and the community regarding division of labor while executing tasks aided by a computer system changes the way a task is conducted and what interaction modalities that are suitable for the overall activity as well as the execution and negotiation of the tasks.

Unfortunately there is not enough space in this short paper to go into our models in more detail, discussing each building block and their relationships with other parts of the model. However, the continuous development of our models based on our cross-disciplinary research proves very promising and we hope to conclude everything in an extensive thesis before the end of the year. Our research on the next generation work support for tactical officers and sensor operators within maritime surveillance, who work closely within highly specialized teams, while making use of different kinds of interaction modalities or combinations of modalities to execute their work, and who negotiate their division of labor, are well suited for our research. We hope that our results will shed some light on the impact of cooperation on the preferred choice of interaction modalities. We believe that our framework will be a welcome help for novice designers of cooperative multimodal systems when making expert decisions in choice of modality or combinations of modalities.

REFERENCES

1. Allwood, J. (2001). Cooperation and Flexibility in Multimodal Communication. Cooperative Multimodal Communication. H. Bunt and R. Beun, Springer Berlin / Heidelberg.
2. Cadier, F. (2007). Modèles Cognitifs pour les Système d'Aide à la Décision Collective: Application à la Patrouille Maritime. École Nationale Supérieur des Télécommunications de Bretagne. PhD.
3. Endsley, M. R. (2000). Situation Awareness Analysis and Measurement. M. R. Endsley and D. J. Garland, Lawrence Erlbaum Associates.
4. Hoogstoel, F. (2001). Les répercussions du travail coopératif assisté par ordinateur sur les systèmes d'information. Environnements évolués et évaluation de l'IHM. C. Kolski. Paris, HERMES Science Europe Ltd.
5. Kuutti, K. (1991). The concept of activity as a basic unit of analysis for CSCW research. Proc. ECSCW'91, Amsterdam, The Netherlands, Kluwer Academic Publishers.
6. Nigay, L. (2004). Design space for multimodal interaction. IFIP Congress Topical Sessions.
7. Nowak, M. (2006). "Five Rules for the Evolution of Cooperation." Science 314(5805): 1560-1563.
8. Vanderdonckt, J., G. Calvary, et al. (2008). Multimodality for Plastic User Interfaces: Models, Methods, and Principles. Multimodal User Interfaces. D. Tzovaras, Springer Berlin Heidelberg: 61-84
9. Verdurand, E. (2011). Modélisation et Évaluation de l'Interaction dans les Système Multimodaux. Télécom Bretagne. PhD
10. Vygotsky, L. S. (1978). Mind in Society: Development of Higher Psychological Processes, Harvard University Press.